Toward Prosodic Contrast: Suai and Pattani Malay

Arthur S. Abramson

Haskins Laboratories, New Haven, Connecticut and the University of Connecticut, Storrs, Connecticut, U.S.A.
abramson@uconnvm.uconn.edu

Abstract

We may have an early stage of tonogenesis in two minority languages of Thailand. In one of them, Suai, its old voice-register distinction is apparently giving way to one of pitch accent. In the other, Pattani Malay, an existing distinction of word-initial consonant length is giving way to one of relative salience on the first syllable of disyllabic words. The role of pitch in both may foretell the ultimate emergence of tones.

1. Introduction

In diachronic phonology an intriguing puzzle has long been the causes of sound change. Phonetic shifts, whatever the causes, may bring about internal structural pressures that alter the sound pattern of a language [13]. In addition, external pressures from another language may initiate such changes [20]. The case of the rise of phonologically distinctive tones from a previously toneless state, tonogenesis [15], has led to much inspired speculation and comparative linguistic reconstruction, e.g., [14], as well as some experimental research (e.g., [7, 11]).

The classical approach to tonogenesis, as exemplified by the work of André-G. Haudricourt [10], ascribes the adding of distinctive pitch contours to syllables as consequences of the coarticulatory effects of properties of initial and final consonants. Specifically, two critical stages are described, at least for Vietnamese and apparently members of other language families in Asia. At the beginning of our Common Era when Vietnamese was still toneless, the language had a voicing contrast in initial stops and three syllable types: (1) syllables ending in a vowel or nasal consonant, (2) syllables ending in /h/, and (3) those ending in /ʔ/. By the end of the 6th century, with the loss of final /h/ and final /ʔ/, two dynamic tones emerged, a falling tone from (2) and a rising tone from (3). A static mid tone arose from the open syllables of (1). In the meantime, we are told, the voicing distinction in initial consonants was weakening. As a result of its ultimate loss, by the end of the 12th century the three Vietnamese tones had been doubled to six because of the enhancement in production and perception of pitch perturbations that had accompanied the releases of the old stop consonants. That is, the pitch at the onset of the vowel at the release of the old voiced stop was lower than that at the release of the old voiceless stops.

This model of the shift from the separation of lexical categories by consonantal properties to tonal ones has been amended by later research and reasoning [19] to include a stage in which the voice qualities (phonation types) of the syllables played a distinctive role. This prompts a summary of recent and current research on two minority languages seemingly undergoing movement toward tonogenesis. In one of them, Suai, phonation type or, more broadly speaking, voice register is the key factor. In the other, Pattani Malay, neither voicing as such nor phonation type bears on the phenomenon. Rather, it is distinctive consonant length in word-initial position.

2. Suai voice registers

Suai is a Mon-Khmer language of Southeast Asia. Many of these languages have distinctive voice registers that include differences in phonation type, pitch, loudness, and other phonetic properties. Normally such registers are named by phonation type, such as clear (modal), breathy, and creaky voice. Phonation type is conventionally taken to be the dominant property [17].

Although acoustic analysis of one of its major dialects, Kui, has been done [18], I know of no research that includes experiments on the perception of Suai voice registers or, for that matter, those of any other such language. Together with colleagues [8] I have been working on the acoustics of the production and perception of the voice registers of Kuai, another major dialect of Suai. Our Kuai speakers live in the village of Samrong in the province of Surin in northeastern Thailand. They also speak tone languages, Thai or Lao or both.

2.1. Production of Kuai registers

Kuai is said to have two voice registers: clear (modal) vs. breathy, as in /luː/ 'to howl' vs. /lṳː/ 'thigh.' Listening to the citation forms of words in minimal or quasi-minimal pairs recorded by six speakers, we were not at all sure that the members of the pairs were always distinguished by phonation type. Sometimes we heard only a pitch difference, sometimes both, sometimes nothing.

This seeming instability in maintaining the distinction in speech is compatible with some uncertainty in the labeling behavior of some of our subjects in the perceptual experiments to be presented below. Acoustic data discussed here will be limited to the two important dimensions of spectral tilt and fundamental frequency (F0).

By spectral tilt I mean the attenuation of higher frequencies of the acoustic spectrum as a consequence of incomplete adduction of the glottis in breathy voice. This is a good acoustic index of breathiness as long as the measurements are taken in steady-state portions of vowels with essentially the same formant frequencies in both members of the paired contrast. Thus, although we had a larger number of word-pairs, for this analysis the foregoing constraints limited us to a choice of eight pairs uttered three times each by our six speakers. The indexes used to measure spectral tilt were the amplitude ratios of the second harmonic to the first (H_2/H_1) and of the harmonic closest to the peak of the first formant and the first harmonic (H_{F1}/H_1). An analysis of variance (ANOVA) revealed that for only three of the speakers, as shown in Table 1, do both the indexes for the difference between the registers reach significance. For the other three speakers neither index was significant.

Table 1: *ANOVA of spectral tilt for three speakers [8].*

Ratios of Harmonics: Speakers A, C and D					
Variable		**H_2/H_1**		**H_{F1}/H_1**	
	df	**F**	**p**	**F**	**p**
Register	1	132.9	.0074	19.14	.0485
Word	7	0.503	.8173	7.746	.0006
Reg x Word	7	0.990	.4765	2.630	.0586

2.2. Kuai Fundamental-frequency contours

We took F0 values at 10 ms intervals from each of the three tokens of each member of the full set of 11 pairs of words distinguished by the two voice registers. To depict the contours as something closer to the sensation of pitch and to facilitate averaging values across our six speakers, three men and three women, we converted Hz to semitones. We normalized the data in time by fitting them all to a common length of 100 arbitrary units. They are plotted in Figure 1.

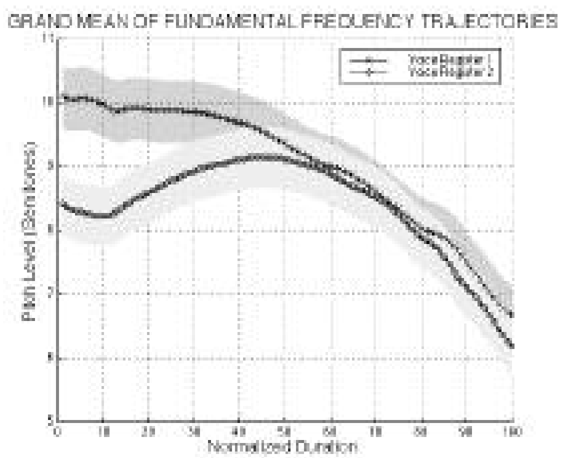

Figure 1: *Kuai F0 contours of six speakers normalized in time and frequency. Register 1 (upper curve) = clear voice; Register 2 = breathy voice. The shaded areas show the standard error of` the means. Adapted from [8].*

.

Upon inspection of the curves in Figure 1, we inferred that the critical difference was in the slopes of the first half of the graph. Accordingly, we divided the time scale into four sections to provide for separate analysis of the portions differing in height and slope: samples 1–15, 16–50, 51–75, and 76–100. Not surprisingly, an ANOVA showed that precisely in section two, where the slopes differ markedly, the difference in gradient is significant. In section 1 the heights are significantly differet.

2.3. Kuai perception of fundamental frequency

Actually, unaware early in our research that the register distinction might no longer be robust, at least in the speech of the village of Samrong, we had prepared an elaborate set of synthetic stimuli for perceptual testing. A control test with the utterances of four native speakers revealed that only eight out of our 16 test subjects responded well above chance to the distinction. For the balance of the perceptual experiments we used only those eight people. Although our acoustic stimuli were designed for an exploratory study in synthesis to test the efficacy of several factors that pertain to voice register [8], namely, open quotient of the simulated glottal period, spectral tilt, amplitude of turbulence, overall relative amplitude, and F0, I am limiting this discussion to the results with stimuli that have F0 as a cue.

To generate stimuli for our experiments we used the SynthWorks program of Scicon Research and Development. It is a commercial version of the synthesizer designed by Dennis Klatt [12]. In the series of 19 synthetic-speech stimuli of concern here F0 was set to begin at 70, 101, and 135 Hz and then reach 120 Hz at 310 ms where it remained for 40 ms and then fell slowly to 99 Hz at the end at 580 ms. The value of 101 Hz at the beginning was part of our default pattern, while the other two values were our manipulations from below and above early in the syllable. Since Suai does not have an established writing system, the subjects identified each item by checking a Thai gloss on the answer sheet meaning either ‘to howl’ or ‘thigh.’ The results can be seen in Figure 2.

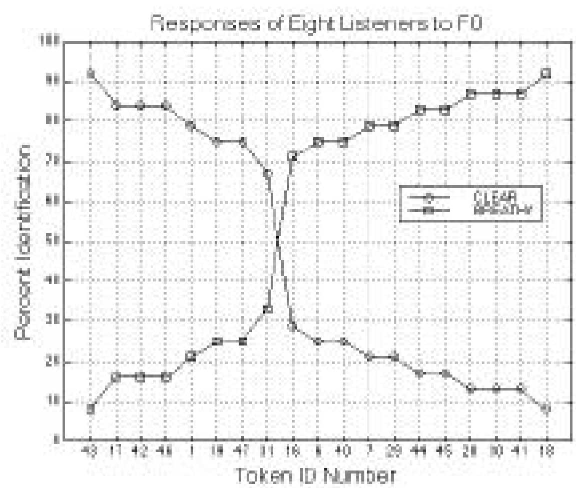

Figure 2. *Percentages of Kuai responses to 19 synthetic syllables with modifications in F0. Adapted from [8].*

The token identification numbers along the bottom of the graph form a key to the parameters manipulated [8]. Most of them include a change on at least one other parameter in addition to F0. Suffice it to say, however, that all the items in the left half cause dominant “clear” responses and, except for number 1, have a high falling F0; number 1 is at the default setting. On the right side after the crossover point of the two functions all the stimuli cause mainly “breathy” responses and, except for numbers 7 and 29, have a low rising F0. In 7 and 29 F0 falls from above; the other parameters of number 7 are at their default values, while number 29 has a large open quotient, which favors breathiness.

3. Pattani Malay consonant length

Pattani Malay is unusual in having distinctive word-initial consonant length as in, e.g., /butɔ/ 'blind' vs. /bːutɔ/ 'a kind of tree,' /pagi/ 'morning' vs. /pːagi/ 'early morning.' All consonantal manners and, except for glottal stop, places of articulation are subject to this distinction. The language is spoken in the southernmost provinces of Thailand just above the border with Malaysia. Its speakers are also fluent in Thai.

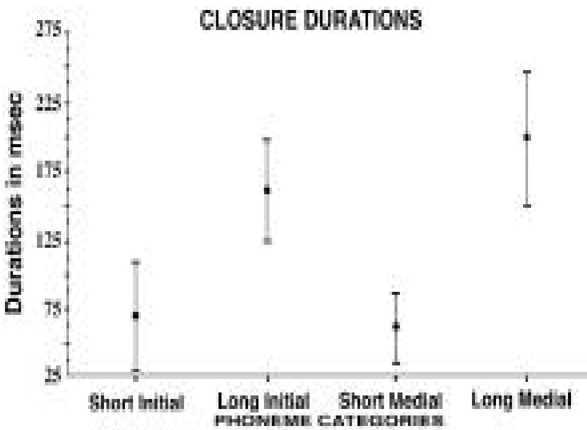

Figure 3. *Means and standard deviations for Pattani Malay speaker PCM. N = 28 for each point. Adapted from [2].*

3.1. Production of the Pattani Malay contrast

In my earliest acoustic approaches to this topic [1, 2] I took the conventional view of distinctive length seriously and measured the durations of consonant closures and, for fricatives, constrictions. For voiceless stops, of course, the silent closures could not be measured in utterance-initial position but rather embedded after a vowel in a short carrier sentence. At the same time, looking ahead to work on perception, I understood that the duration as such of utterance-initial silent closures would be inaudible. Given the finding in the meantime that voiceless stops in utterance-initial position were indeed well identified, I also measured relative amplitudes of the typically disyllabic words with the hypothesis that for aerodynamic reasons the first syllable of a word with a long initial voiceless stop would be more salient. Later [4] it became expedient also to measure fundamental frequency (F0).

In Figure 3 are shown the means and standard deviations of the closure durations for one speaker in both initial and medial position [2]. As explained above, the voiceless stops, as well as voiceless affricates, are included only in the medial data. An ANOVA showed duration to be highly significant in both contexts.

Examining four speakers' utterances of citation forms of words with short and long initial voiceless stops, voiced stops, and continuants, I obtained the ratios of the peak amplitudes of the first syllable to the second [4]. On average, the ratio was three times greater for words with initial long consonants. Paired *t*-tests showed the results to be highly significant for voiceless stops and moderately significant for voiced stops and continuants.

3.2. Perception of the Pattani Malay contrast

The data from the aforementioned control tests with citation forms of words in natural speech are given in Figure 4. Recognition rates are high, even for voiceless stops. The labeling of voiceless affricates, however, is a bit above chance.

Through waveform editing, in a series of experiments I incrementally lengthened the closures of short consonants and shortened the closures of long consonants in natural speech, while leaving everything else in the digitized speech signals untouched [1]. The results were unambiguous. As duration increased, the perception of original short consonants shifted to long; as duration decreased, the reverse happened. Different crossover points in the labeling functions indicated the presence of at least one more cue.

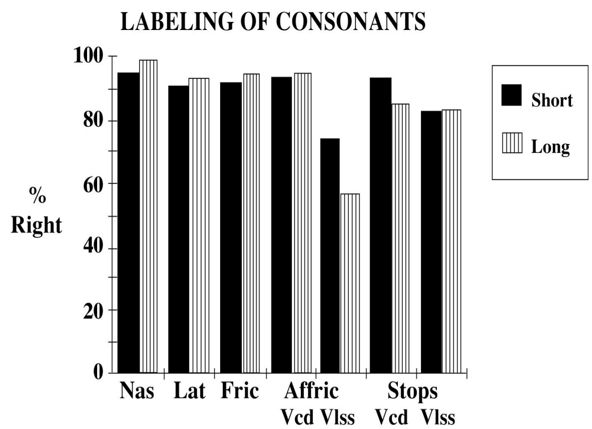

Figure 4. *Identifications by 21 subjects of 32 randomized Pattani Malay words recorded by two speakers [1].*

Experiments with manipulation in natural speech of amplitude ratios [3] and F0 ratios [5] yielded significant changes in perception; nevertheless, these changes were not extensive enough to effect complete crossovers in the labeling functions. That is, neither of these factors served alone as a sufficient cue.

Bearing in mind that it is probably the rare phonological distinction that has underlying it only a single phonetic property, I performed perceptual experiments with computer-controlled co-variation of both F0 and amplitude [6]. The results of one of the tests, the one with original short /k/ from /kitɔ/ 'we, us' (contrasted with /kːitɔ/ 'among us'), are displayed in Figure 5. As for Figure 2, so also for Figure 5, there is not enough space to give a table of parameter values for the stimuli. Also, Figure 5 is necessarily small in the space available. Perhaps a brief explanation and a few examples from the two ends of the series will help.

For the first syllable of a word with a short initial consonant F0 was increased in steps of 0, 2, and 3 semitones and amplitude in steps of 0, 3, and 6 dB. The second syllable was decreased in F0 in steps of 0, -0.75, and -1.5 ST and 0, -3, and –6 dB. At the far left of the graph in Figure 5 k1 is the unchanged original utterance. Second from the left, stimulus k16 has a 3 dB rise in the first syllable and no other change. At the far right end k33 has a rise of 3 semitones and 6 dB on the first syllable but only a 6 dB fall on the second syllable. Left of it. stimulus k31 likewise has +3 ST and +6dB on the first syllable but only -1.5 ST on the second syllable.

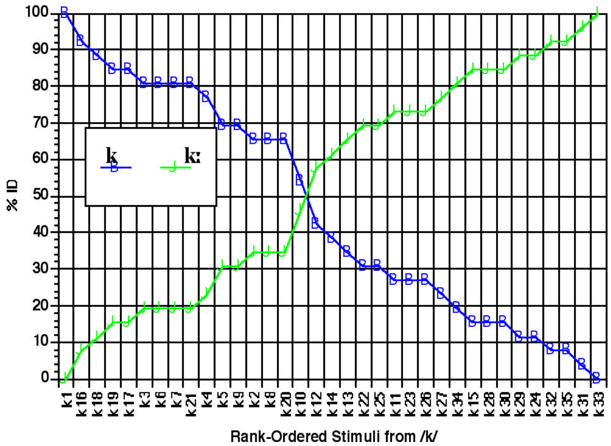

Figure 5. *Responses of 26 Pattani Malay subjects to stimuli derived from short /k/. Labels heard as short start from the upper left; those heard as long rise to the right. Adapted from [6].*

4. Conclusions

Experimental findings and impressionistic observations imply that both languages, Suai and Pattani Malay, are pursuing different paths leading to phonological shifts from clear and breathy voice registers for the former, and for the latter, from word-initial distinctive consonant length, to a kind of prosodic salience. This could be a matter of a replacement by phonemic stress or accent, yet, given the close contact with Thai, a tone language, and the widespread bilingualism of the speakers of the two minority languages, we may have here a way station on the road to tonogenesis. One source [16] claims that it has already happened in Suai. I am not convinced. Another [9] reconstructs such a role for phonation type in a different branch of Mon-Khmer.

5. Acknowledgment

This research was supported by NIH Grant DC-02717 to Haskins Laboratories and a grant from the Thailand Research Fund. I am also grateful to Dr. Theraphan Luangthongkum and her colleagues at Chulalongkorn University, Bangkok for their help and warm hospitality.

6. References

[1] Abramson, A.S., 1986. The perception of word-initial consonant length: Pattani Malay. *Journal of the International Phonetic Association, 16,* 8–16.

[2] Abramson, A. S., 1987. Word-initial consonant length in Pattani Malay. *Proceedings of the 11th International Congress of Phonetic Sciences, Vol. 6*. Tallinn: Academy of Sciences of the Estonian S.S.R., 68–70.

[3] Abramson, A.S., 1991. Amplitude as a cue to word-initial consonant length: Pattani Malay. *Proceedings of the 12th International Congress of Phonetic Sciences*, Vol. 3. Aix-en-Provence: Université de Provence, 98–101.

[4] Abramson, A. S., 1998. The complex acoustic output of a single articulatory gesture: Pattani Malay word-initial consonant length. In *Papers from the 4th Annual Meeting of the Southeast Asian Linguistics. Society 1994.* U. Warotamasikkhadit & T. Panakul (eds.). Tempe, Arizona: Arizona State University Program for Southeast Asian Studies, Monograph Series Press, 1–20.

[5] Abramson, A. S., 1999. Fundamental frequency as a cue to word-initial consonant length: Pattani Malay. *Proceedings of the 14th International Congress of Phonetic Sciences.* Berkeley, CA: University of California Press, 591–594.

[6] Abramson, A.S., 2003. Acoustic cues to word-initial stop length in Pattani Malay. *Proceedings of the 15th International Congress of Phonetic Science, Barcelona, 3-9 August 2003.* Barcelona: The 15th ICPhS Organizing Committee, 387–390.

[7] Abramson, A.S.; Erickson, D.M., 1992. Tone splits and voicing shifts in Thai: Phonetic plausibility. In *Pan-Asiatic Linguistics: Proceedings of the Third International Symposium on Languages and Linguistics, Vol. I .* Bangkok: Chulalongkorn University, 1–16.

[8] Abramson, A.S., L-Thongkum, T., Nye, P., In preparation. The acoustics of the production and perception of voice registers in Suai.

[9] Diffloth, G., 1982. Registres, dévoisement, timbres vocaliques: Leur histoire en Katouique. *Mon-Khmer Studies* 11, 47–82.

[10] Haudricourt, A-G., 1954. De l'origine des tons en vietnamien. *Journal asiatique* 242, 69–82

[11] Hombert, J-M., Uhala, J.J., Ewan. W.G., 1979. Phonetic explanations for the development of tones. *Language* 55, 37–58.

[12] Klatt, D.H. 1980. Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America* 67, 971–995.

[13] Martinet, A., 1955. *Économie des changements phonétiques: Traité de phonologie diachronique.* Berne: Éditions A, Francke S.A.

[14] Maspero, H., 1911. Contribution à l'étude du système phonétique des langues thai. *Bulletin de l'École française d'ExtrêmeOrient* 11, 153–169.

[15] Matisoff. J.A., 1973. Tonogenesis in Southeast Asia. In *Consonant Types & Tone.* L.M. Heyman (ed.). Southern California Occasional Papers in Linguistics, No.1. Los Angeles: Linguistics Program, U. of Southern California, 71–85.

[16] Sukkasame, P., In press. Tonal evolution in Suai (Kuay). *Papers from the Eleventh Annual Meeting of the Southeast Asian Linguistics Society, 2001.*

[17] Thongkum, T.L., 1988. Phonation types in Mon-Khmer languages. In *Vocal Physiology: Voice Production, Mechanisms and Functions.* O. Fujimura (ed.). New York: Raven Press, 319–333.

[18] Thongkum, T.L., 1989. An acoustic study of the register comples in Kui (Suai). *Mon-Khmer Studies,* 15, 1–19.

[19] Thurgood, G., 2000. Vietnamese and tonogenesis: Revising the model and the analysis. *Diachronica* 19, 333–363.

[20] .Weinreich, U., 1953. *Languages in contact.* New York: Linguistic Circle of New York.